Estimation of the Mean of a Stationary Random Process by Periodic Sampling

By H. T. BALCH, J. C. DALE, T. W. EDDY and R. M. LAUVER

(Manuscript received February 23, 1966)

Estimating the mean of a stationary random process from the average of equally weighted samples taken periodically in a closed interval (0, T) is considered. The variance of this estimator as a function of the number of samples in the interval is given in the form of a modified sampling theorem.

I. INTRODUCTION

This paper* considers the problem, commonly encountered in detection theory, of estimating the mean of a stationary random process from samples taken periodically in a closed interval $(0, T)$. The samples are, in general, correlated and the estimator used is the average of equally weighted samples taken in the interval. Existing results are extended to give a clearer interpretation of the dependence of the variance on the number of samples. The dependence is obtained in terms of the power spectral density of the process in the form of a modified sampling theorem.

II. THEORY

2.1 *General*

To estimate the mean value, A, of $s(t)$ where

$$s(t) = A + n(t), \tag{1}$$

the first sample is taken at $t = 0$, and a total of $N + 1$ samples is taken in time T. $n(t)$ is a sample function from a wide-sense stationary random process with mean zero and known autocorrelation function $R(\tau)$.

* This work was supported by the U. S. Navy, Bureau of Ships under contract No. Nobsr-89401.

The estimator of A is

$$\hat{A} = [1/(N+1)] \sum_{m=0}^{N} s(mT/N). \tag{2}$$

For a fixed T, N is to be chosen to minimize the variance of $\hat{A}$.

The variance of $\hat{A}$ is given by

$$\sigma^2(\hat{A}) = [1/(N+1)] \sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) R(mT/N). \tag{3}$$

Equation (3) may be found in slightly different form in the literature.[1,2]

It is now convenient to define a weighting function, $q_{\tau_o}(\tau)$, by

$$q_{\tau_o}(\tau) = \begin{cases} 1 - \dfrac{|\tau|}{\tau_o} & |\tau| \leqq \tau_o \\ 0 \quad , & \text{otherwise.} \end{cases} \tag{4}$$

With this definition, (3) may be written as

$$\sigma^2(\hat{A}) = \frac{1}{(N+1)} \int_{-\infty}^{\infty} q_{[(N+1)/N]T}(\tau) R(\tau) \sum_{m=-\infty}^{\infty} \delta\left(\tau - \frac{mT}{N}\right) d\tau, \tag{5}$$

where $\delta(\tau)$ is the Dirac delta function. It is more revealing to express the variance in terms of spectral densities; thus, we make the following definitions:

$$\begin{aligned} F(\omega) &= \frac{1}{(N+1)} \int_{-\infty}^{\infty} q_{[(N+1)/N]T}(\tau) R(\tau) \exp(-j\omega\tau)\, d\tau \\ &= Q(\omega) \circledast S(\omega), \end{aligned} \tag{6}$$

where

$$\begin{aligned} Q(\omega) &= \frac{1}{(N+1)} \int_{-\infty}^{\infty} q_{[(N+1)/N]T}(\tau) \exp(-j\omega\tau)\, d\tau \\ &= \frac{T}{N} \left\{ \frac{\sin [\omega(N+1)/2N]T}{[\omega(N+1)/2N]T} \right\}^2 \end{aligned} \tag{7}$$

and $S(\omega)$ is the spectral density of $n(t)$.

Also define

$$\begin{aligned} G(\omega) = \frac{1}{(N+1)} \int_{-\infty}^{\infty} & q_{[(N+1)/N]T}(\tau) R(\tau) \\ & \exp(-j\omega\tau) \sum_{m=-\infty}^{\infty} \delta\left(\tau - \frac{mT}{N}\right) d\tau. \end{aligned} \tag{8}$$

By using Poisson's sum formula we can show that

$$G(\omega) = \frac{N}{T} \sum_{m=-\infty}^{\infty} F\left(\omega - m \frac{2\pi N}{T}\right). \tag{9}$$

Now, comparing (5), (8), and (9), we observe that

$$\sigma^2(\hat{A}) = G(0) = \frac{N}{T} \sum_{m=-\infty}^{\infty} F\left(m \frac{2\pi N}{T}\right). \tag{10}$$

Because $Q(\omega)$ is approximately zero for

$$|\omega| \geqq (2\pi/T)[N/(N+1)],$$

if $S(\omega)$ is zero for $|\omega| \geqq 2\pi B$, their convolution, $F(\omega)$, will be approximately zero for $|\omega| \leqq 2\pi[B + N/T(N+1)]$. From this result and (10) we observe that choosing

$$\frac{2\pi N}{T} \geqq 2\pi \left[B + \frac{1}{T} \frac{N}{(N+1)}\right] \tag{11}$$

makes

$$\sigma^2(\hat{A}) = G(0) \approx (N/T)F(0). \tag{12}$$

Although the restriction of (11) appears to minimize the variance of $\hat{A}$, it should be observed that $F(0)$ is also a function of N, namely

$$\frac{N}{T} F(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \left\{ \frac{\sin\left[\omega \dfrac{(N+1)}{2N} T\right]}{\omega \dfrac{(N+1)}{2N} T} \right\}^2 d\omega. \tag{13}$$

If $2\pi/T$ is of the same order of magnitude as $2\pi B$, the bandwidth of $S(\omega)$, then $(N/T)F(0)$ and therefore the $\sigma^2(\hat{A})$ may be minimized by choosing the smallest value of N satisfying (11). In such cases, making N larger may actually increase the variance, as illustrated in the examples. $(N/T)F(0)$ would be independent of N if the first sample is taken at $t = T/N$ rather than $t = 0$. Solving (11) for N yields an approximate rule that

$$N \approx BT \frac{1 + \sqrt{1 + (4/BT)}}{2}, \tag{14}$$

where N is an integer, will minimize the variance of $\hat{A}$. It should be recalled that the total number of samples taken in T is $N + 1$.

The form of (9) is frequently encountered in sampling theory, where one sometimes thinks in terms of the original spectra, $F(\omega)$, shifted

by integral multiples of the sampling frequency, $2\pi N/T$. Tbe Nyquist frequency is determined such that overlapping of the sideband spectra is small. This is too restrictive when one is interested in the variance, since then only the value of $G(\omega)$ at $\omega = 0$ is of interest. Thus, sampling can be done at a rate sufficient to prevent overlapping of sidebands at $\omega = 0$. Equation (14) may be considered as a modified sampling theorem, stating that, to minimize the variance, the sampling frequency, f_s, must satisfy

$$f_s = \frac{N}{T} = B\,\frac{1 + \sqrt{1 + (4/BT)}}{2}\,. \tag{15}$$

For large T, f_S is equal to one half of the Nyquist frequency required to reconstruct the time function.

2.2 *Variance for Large T*

When T is sufficiently large, the $Q(\omega)$ function approaches a delta-function, namely

$$Q(\omega) \approx 2\pi\delta(\omega)/(N + 1), \tag{16}$$

and

$$\sigma^2(\hat{A}) \approx [N/T(N + 1)] \sum_{m=-\infty}^{\infty} S(m2\pi N/T). \tag{17}$$

If, as before, $S(\omega)$ is bandlimited and

$$2\pi N/T \geqq 2\pi\left[B + \frac{1}{T}\,\frac{N}{(N + 1)}\right] \approx 2\pi B \tag{18}$$

then

$$\sigma^2(\hat{A}) \approx S(0)/T. \tag{19}$$

Notice that taking more than BT samples will not decrease the variance appreciably. Taking less than BT samples will increase the variance at a rate which depends on $S(\omega)$. The dependence of the variance on N can be easily obtained for this limiting situation from (17). In general, if $dS(\omega)/d\omega \leqq 0$ for $\omega \geqq 0$, then the variance of $\hat{A}$ will also be a monotonically decreasing function of N for $N/(N + 1) \approx 1$. On the other band, if the spectral density of the noise is not monotonically decreasing for $\omega \geqq 0$, then the variance of $\hat{A}$ will have local minima for values of $N < BT$. These statements concerning monotonicity would be true for all T and N if $(N/T)F(0)$ were independent of N.

2.3 *Limit of Continuous Sampling*

The limit of continuous sampling has been derived elsewhere[1] and is easily obtained from (3). The result is

$$\sigma^2(\hat{A})_C = \frac{2}{T}\int_0^T \left(1 - \frac{\tau}{T}\right) R(\tau)\, d\tau = \frac{1}{T}\int_{-\infty}^{\infty} q_T(\tau)R(\tau)\, d\tau \quad (20)$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)\left(\frac{\sin \omega T/2}{\omega T/2}\right)^2 d\omega. \quad (21)$$

As T becomes large,

$$\sigma^2(\hat{A})_C \approx \int_{-\infty}^{\infty} \frac{S(\omega)}{T}\, \delta(\omega) d\omega = \frac{S(0)}{T} \quad (22)$$

Thus, for large T, taking $N = BT$ samples gives the same variance as sampling continuously.

2.4 *Equivalent Independent Samples*

When T is large, one can determine the number of independent samples required to achieve the same variance as continuous sampling. The variance for N_i independent samples is

$$\sigma^2(A)_{N_i} = \frac{\dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} S(\omega)d\omega}{N_i}. \quad (23)$$

Equating this variance to the variance of (22) requires

$$N_i = T\,\frac{\dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} S(\omega)d\omega}{S(0)}. \quad (24)$$

Defining the effective bandwidth as

$$2B_e = \frac{\dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} S(\omega)d\omega}{S(0)}, \quad (25)$$

(24) can be written as

$$N_i = 2B_e T. \quad (26)$$

However, the variance achieved by continuous sampling may be obtained by taking $N = BT$ samples in time T. Thus,

$$N_i = (2B_e/B) \quad (27)$$

equates N for minimum variance to the number of independent samples required to achieve the same variance.

III. EXAMPLES

The variance of the sample mean as a function of number of samples ($N + 1$) and length of record (T) has been computed for several spectral densities.

The variance of the sample mean shown on the following figures was computed using (3).

3.1 *Rectangular Spectrum*

$$S_1(\omega) = \begin{cases} \frac{1}{2}, & -2\pi < \omega < 2\pi \\ 0, & \text{elsewhere.} \end{cases} \tag{28}$$

Fig. 1 shows $\sigma^2(\hat{A})$ plotted against number of samples. Each curve of the set represents a different length of record T. The table on the figure shows the relationship of the curves to the length of record.

The most striking feature of the curves on Fig. 1 is the abrupt steps

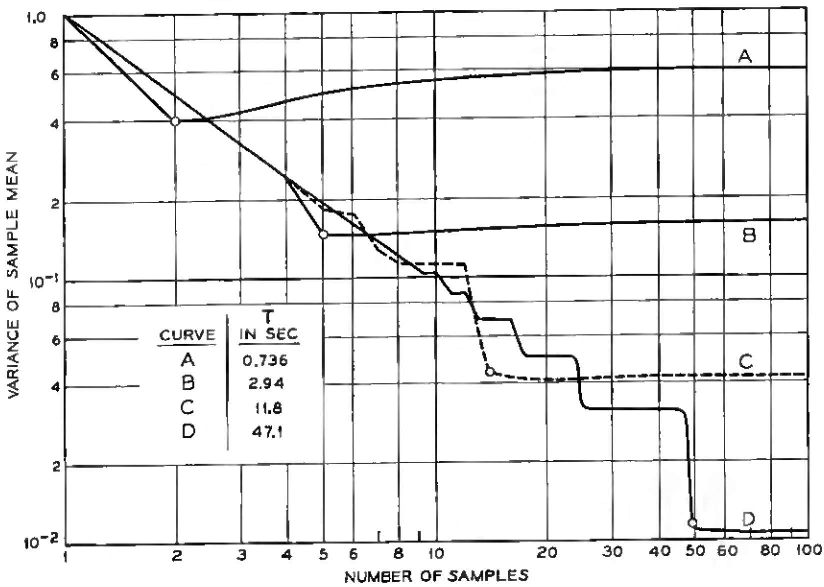

Fig. 1 — Variance of the sample mean as a function of the number of samples and length of record for a process with rectangular spectral density.

in $\sigma^2(\hat{A})$ as N is increased. This behavior is predicted by (17). Equation (14) predicts the approximate value of N for minimum variance. These values of N are shown for each of the curves by a small circle.

An interesting point to note here is that in some regions a better estimate of the mean is obtained when the same number of samples is taken for a smaller T. Also for small T, $\sigma^2(\hat{A})$ reaches a minimum and then increases as more samples are taken. This implies that for small values of T a smaller variance is obtained by taking a smaller number of samples (but including the end points) than would be obtained by continuous sampling.

3.2 *Sawtooth Spectrum*

$$S_2(\omega) = \begin{cases} \left| \dfrac{\omega}{2\pi} \right|, & -2\pi \leqq \omega \leqq 2\pi \\ 0\ , & \text{elsewhere.} \end{cases} \tag{29}$$

This is an interesting case for two reasons. First, its spectrum is not monotonically decreasing. This gives rise to local minima and maxima in $\sigma^2(\hat{A})$ as a function of N caused by the spectrum shape.

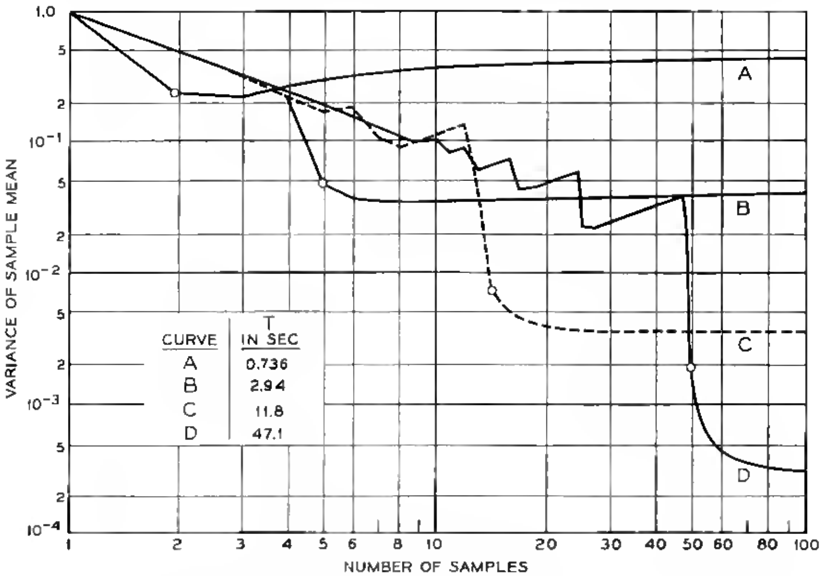

Fig. 2 — Variance of sample mean as a function of number of samples and length of record for a process with a sawtooth spectral density.

Second, its spectrum at $\omega = 0$ is 0, thus enhancing the error due to approximating $Q(\omega)$. The results are shown in Fig. 2.

3.3 *Markoff Spectrum*

$$S_3(\omega) = \frac{8}{\omega^2 + 16}. \tag{30}$$

This is an example of a nonbandlimited spectrum. The values of $\sigma^2(\hat{A})$ are shown in Fig. 3. A point worth noting here is that if the bandwidth of the process was defined as the width at the one-half power points and the time function sampled according to (14), the value of $\sigma^2(\hat{A})$ obtained would be larger by about a factor of 2 than the minimum value obtained by letting N approach infinity.

This example is also the same one treated by Fine and Johnson[3] for small values of T. Curve A on Fig. 3 agrees with their results.

IV. SUMMARY

Theory has been presented which predicts the behavior of the variance of the sample mean of periodic samples taken from a stationary random

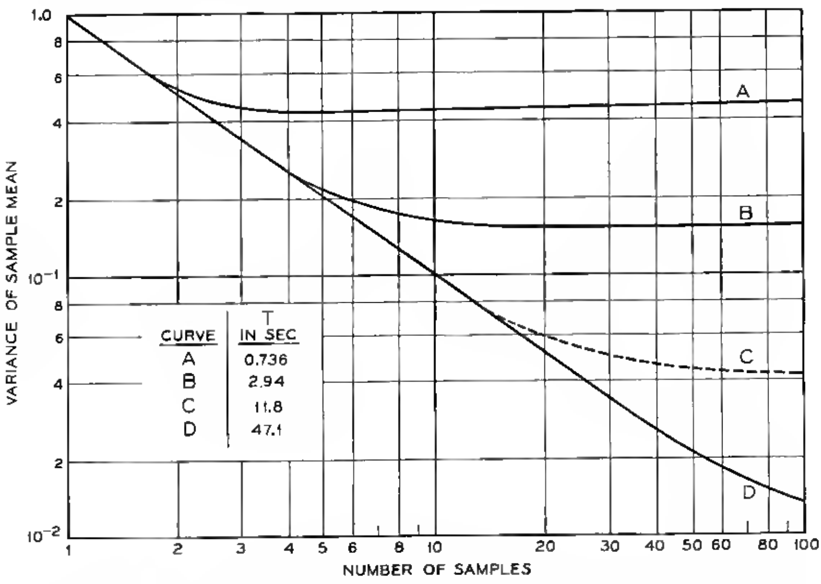

Fig. 3 — Variance of the sample mean as a function of the number of samples and length of record for a process with Markoff spectral density.

process. The variance is given in terms of the power spectrum of the sampled process. Three interesting results have been shown:

(*i*) When $BT \gg 1$, the variance of the sample mean is essentially minimized when BT samples are taken.

(*ii*) The variance of the sample mean is not necessarily monotonically decreasing as a function of the number of samples taken in a fixed record length.

(*iii*) For short record lengths, it is possible to obtain a smaller variance with a small, finite number of samples than with continuous sampling.

REFERENCES

1. Davenport, W. B. Jr. and Root, W. L., *Random Signals and Noise,* McGraw-Hill Book Company, New York, 1958.
2. Costas, J. P., Periodic Sampling of Stationary Time Series, MIT Technical Report No. 156, 16 May, 1950, p. 4, (11).
3. Fine, T. and Johnson, N., On the Estimation of the Mean of a Random Process, Proc. IEEE, Feb., 1965, pp. 187–188.